Stochastic Calculus, Courant Institute, Fall 2019
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2019/StochCalc.html
Jonathan Goodman, September, 2019

Class 12, Ito, theory

1 Introduction

Last week was the setup for Ito theory. This week is the theory. See the Class 11 notes for notation and motivation.

The argument here is different from the arguments in other places. If you look in Wikipedia or other standard sources, you will find a slick looking proof based on "simple functions" and the Ito isometry formula. The details of this approach are not as simple.

2 The Ito integral

The time step is $\Delta t_m = 2^{-m}$. This has two features that make our proof work. One is that $\Delta t_m \to 0$ exponentially fast, which makes the sums in Borel Cantelli converge. The other is that $\Delta t_{m+1} = \frac{1}{2}\Delta t_m$. This makes it easy to compare the Δt_m approximation to the $\Delta t_{m+1} = \frac{1}{2}\Delta t_m$ approximation.

The discrete times are $t_j = j\Delta t$. The approximation to

$$X_t = \int_0^t f_s dW_s \; .$$

is

$$X_t^{(m)} = \sum_{t_j < t} f_{t_j} \left(W_{t_{j+1}} - W_{t_j} \right) \; . \tag{1}$$

We will show that for almost every Brownian motion path $W_{[0,t]}$, the limit $\lim_{m\to\infty}$ exists. This will be our definition of the Ito integral with respect to Brownian motion.

Our proof requires quantitative assumptions on the continuity of the integrand f_s. We are giving a "quantitative" proof based on specific inequalities being used to show a Borel Cantelli sum is finite. We will assume that f_s is "about as continuous" as Brownian motion. Last week we had the specific formula involving the Ito integral

$$X_t = \int_0^t a(X_s)ds + \int_0^t b(X_s)dW_s \; .$$

Here, X_s is a diffusion that is "about as continuous" as Brownian motion. Brownian motion, informally, has $\Delta W \sim \Delta t^{\frac{1}{2}}$, or $(\Delta W)^2 \sim \Delta t$. We assume that there is a constant C (for "continuity") so that

$$\mathrm{E}\Big[\left(f_{s+\Delta t} - f_s\right)^2 \Big] \le C\Delta t \; . \tag{2}$$

This should be true about $b(X_s)$ because in a time increment Δt, we should have $\Delta b \sim b'(X_s)\Delta X$.

The calculation compares $X_t^{(m+1)}$ to $X_t^{(m)}$. More specifically, we find an inequality (an "upper bound") of the form

$$\mathrm{E}\left[\left(X_t^{(m+1)} - X_t^{(m)}\right)^2\right] \leq AC\Delta t_m \;. \tag{3}$$

Here, C is the continuity constant from (2) and A is another constant. For the Borel Cantelli lemma, we need the expected value of $\left|X^{(m+1)} - X^{(m)}\right|$, not the square of this that appears in (3). This may be done using the *Cauchy Schwarz* inequality (which we described in an earlier class?) If U and V are any two random variables with any joint distribution, then

$$\mathrm{E}[UV] \leq \left(\mathrm{E}\left[U^2\right]\mathrm{E}\left[V^2\right]\right)^{\frac{1}{2}} \;. \tag{4}$$

We apply this with $U = \left|X^{(m+1)} - X^{(m)}\right|$ and $V = 1$, so $U^2 = \left(X^{(m+1)} - X^{(m)}\right)^2$ and $\mathrm{E}\left[V^2\right] = 1$. The result is the bound of the expected value in terms of the expected square

$$\mathrm{E}[U] \leq \left(\mathrm{E}\left[U^2\right]\right)^{\frac{1}{2}} \;. \tag{5}$$

[There is more about this inequality in a later section of this class.] We use this and the bound (3). In the end we use the geometric series definition $\Delta t_m = 2^{-m}$. This implies that even $\sqrt{\Delta t_m}$ is a geometric series. Specifically, $\sqrt{\Delta t_m} = z^m$ with $z = 2^{-\frac{1}{2}} < 1$. The calculation is

$$\begin{aligned}
\mathrm{E}\left[\left|X_t^{(m+1)} - X_t^{(m)}\right|\right] &\leq \left(\mathrm{E}\left[\left(X_t^{(m+1)} - X_t^{(m)}\right)^2\right]\right)^{\frac{1}{2}} \\
&\leq (AC\Delta t_m)^{\frac{1}{2}} \\
\mathrm{E}\left[\left|X_t^{(m+1)} - X_t^{(m)}\right|\right] &\leq Bz^m \;,\quad z = 2^{-\frac{1}{2}} \;,\quad B = \sqrt{AC} \;. \qquad (6)
\end{aligned}$$

Thus, the expected square inequality (3) is enough for the Borel Cantelli lemma of last week to show that the limit $m \to \infty$ of the approximate Ito integrals (1) exists. This means that we prove the Ito integral exists by verifying the inequality (3).

You might wish for a more direct approach to our target quantity $\left|X_t^{(m+1)} - X_t^{(m)}\right|$. Is the square $\left(X_t^{(m+1)} - X_t^{(m)}\right)^2$ plus Cauchy Schwarz a roundabout way to do a simple thing? For one thing, the relation between the quantities (5) is (see later today) central to much theory of probability, so we don't think of it as complicated. Also, as we are about to see, computing the square is a way to capture cancellations that arise from the fact that (we will soon see) we are summing a collection of uncorrelated random variables with expected value equal to zero.

To illustrate this point, suppose U_j is a sequence of uncorrelated random variables with mean zero and variance σ^2. These do not need to be independent. The sum is $S_n = \sum_{j=1}^n U_j$. Each term in the sum is on the order of 1, so you might expect the sum to of order n. The triangle inequality gives a bound like this:

$$\mathrm{E}[|S|] = \mathrm{E}[\,|U_1 + \cdots + U_n|] \leq \mathrm{E}[\,|U_1|] + \cdots + \mathrm{E}[\,|U_1|] = n\mathrm{E}[\,|U_1|] = O(n) \; .$$

This is true even if $\mathrm{E}[U_1] \neq 0$. But, if $\mathrm{E}[U_1] = 0$, then the positive and negative terms in the sum roughly cancel. The result is that the sum is likely to be much smaller than $O(n)$. This cancellation doesn't require the terms to be independent, but if they are equal then there is no cancellation. We suppose they are uncorrelated, which is $\mathrm{E}[U_j U_k] = 0$ if $j \neq k$. We will calculate that $\mathrm{E}\big[S^2\big] = O(n)$. The bound (5) then gives $\mathrm{E}[|S|] = O(n^{\frac{1}{2}})$. This is smaller than $O(n)$. The absolute value $|S|$ is not algebraic, but the square S^2 is algebraic. It can be expanded as a sum to find cancellation.

The actual calculation uses the familiar trick that the square of a sum is a double sum.

$$S^2 = \sum_{j=1}^n \sum_{k=1}^n U_j U_k \; .$$

Therefore

$$\mathrm{E}\big[\,S^2\big] = \sum_{j=1}^n \sum_{k=1}^n \mathrm{E}[\,U_j U_k] \; .$$

By assumption, the off diagonal ($j \neq k$) terms are zero. The $j = k$ terms that remain give

$$\mathrm{E}\big[\,S^2\big] = \sum_{j=1}^n \mathrm{E}\big[\,U_j^2\big] = n\sigma^2 \; .$$

This shows that $\mathrm{E}[|S|] \leq \sigma\sqrt{n}$, as claimed. The calculation may be familiar from the central limit theorem. But here we have not assumed that the U_j are independent, only uncorrelated.

Finally, we compare $X^{(m)}$ to $X^{(2m)}$. The discrete times for the two approximations are

$$t_j^{(m)} = j\Delta t_m \; , \;\; t_j^{(m+1)} = j\left(\tfrac{1}{2}\Delta t_m\right) \; .$$

We simplify the notation and calculations by dropping the m and t. We write

$$t_j \text{ instead of } t_j^{(m)} = j\Delta t = j2^{-m} \; , \;\; \Delta t \text{ instead of } \Delta t_m = 2^{-m} \; .$$

The $m+1$ discrete times are related to the m times by

$$\begin{aligned} t_{2j}^{(m+1)} &= t_j^{(m)} = j\Delta t \\ t_{2j+1}^{(m+1)} &= t_j^{(m)} + \tfrac{1}{2}\Delta t \\ &= \left(j + \tfrac{1}{2}\right)\Delta t \; . \end{aligned}$$

The last line suggests the notation $t_{j+\frac{1}{2}}$ for $t^{(m+1)}_{2j+1}$. Finally, we write f_j for f_{t_j} and $f_{j+\frac{1}{2}}$ for $f_{t_{j+\frac{1}{2}}}$. We do this for W also.

One term in the sum (1) defining $X^{(m)}$ is

$$f_j \left(W_{j+1} - W_j\right) \ .$$

In the $X^{(m+1)}$ approximation, this is replaced by two terms

$$f_j \left(W_{j+\frac{1}{2}} - W_j\right) + f_{j+\frac{1}{2}} \left(W_{j+1} - W_{j+\frac{1}{2}}\right) \ .$$

The difference is

$$U_j = f_j \left(W_{j+\frac{1}{2}} - W_j\right) + f_{j+\frac{1}{2}} \left(W_{j+1} - W_{j+\frac{1}{2}}\right) - f_j \left(W_{j+1} - W_j\right) \ .$$

This calculation is the basis of everything:

$$\begin{aligned} U_j &= f_j W_{j+\frac{1}{2}} - f_j W_j + f_{j+\frac{1}{2}} W_{j+1} - f_{j+\frac{1}{2}} W_{j+\frac{1}{2}} - f_j W_{j+1} + f_j W_j \\ &= f_{j+\frac{1}{2}} \left(W_{j+1} - W_{j+\frac{1}{2}}\right) f_j \left(W_{j+1} - W_{j+\frac{1}{2}}\right) \\ U_j &= \left(f_{j+\frac{1}{2}} - f_j\right) \left(W_{j+1} - W_{j+\frac{1}{2}}\right) \ . \end{aligned} \tag{7}$$

The difference we are trying to estimate has the form

$$X^{(m+1)}_t - X^{(m)}_t = S = \sum_{t_j < t} U_j \ .$$

The bound (3) is a bound on $\mathrm{E}[S^2]$, with its diagonal and off diagonal terms.

We use the tower property, conditioning on information available up to the time of the start of the last increment of Brownian motion, which is $t_{j+\frac{1}{2}}$. Let Q be anything random, and R something known at time $t_{j+\frac{1}{2}}$, which means $R = R(W_{[0,t_{j+\frac{1}{2}}]})$, then

$$\begin{aligned} \mathrm{E}[QR] &= \mathrm{E}\Big[\, \mathrm{E}\Big[\, QR(W_{[0,t_{j+\frac{1}{2}}]}) \mid \mathcal{F}_{j+\frac{1}{2}} \Big]\Big] \\ &= \mathrm{E}\Big[\, \mathrm{E}\Big[\, Q \mid \mathcal{F}_{j+\frac{1}{2}} \Big]\, R(W_{[0,t_{j+\frac{1}{2}}]})\Big] \end{aligned}$$

For the off diagonal terms $U_j U_k$ with $j \neq k$, we assume that $j > k$ and take $Q = W_{j+1} - W_{j+\frac{1}{2}}$. The rest of $U_j U_k$ is

$$R = \left(f_{j+\frac{1}{2}} - f_j\right) \left(f_{k+\frac{1}{2}} - f_k\right) \left(W_{k+1} - W_{k+\frac{1}{2}}\right) \ .$$

All of this is known at time $t_{j+\frac{1}{2}}$. Therefore,

$$\mathrm{E}[\, U_j U_k] = \mathrm{E}\Big[\, \mathrm{E}\Big[\, W_{j+1} - W_{j+\frac{1}{2}} \mid \mathcal{F}_{j+\frac{1}{2}} \Big] \left(f_{k+\frac{1}{2}} - f_k\right) \left(W_{k+1} - W_{k+\frac{1}{2}}\right)\Big] \ .$$

This is equal to zero because the inner expectation has an increment of Brownian motion in the future of $\mathcal{F}_{j+\frac{1}{2}}$. The independent increments property makes $W_{j+1} - W_{j+\frac{1}{2}}$ independent of anything known at time $t_{j+\frac{1}{2}}$.

The diagonal terms may be calculated using similar reasoning:

$$\mathrm{E}\left[\, U_j^2 \right] = \mathrm{E}\left[\mathrm{E}\left[\left(W_{j+1} - W_{j+\frac{1}{2}} \right)^2 \mid \mathcal{F}_{j+\frac{1}{2}} \right] \left(f_{j+\frac{1}{2}} - f_j \right)^2 \right] \,.$$

The independent increments property allows us to calculate

$$\mathrm{E}\left[\left(W_{j+1} - W_{j+\frac{1}{2}} \right)^2 \mid \mathcal{F}_{j+\frac{1}{2}} \right] = \tfrac{1}{2}\Delta t \,.$$

This is because $\left(W_{j+1} - W_{j+\frac{1}{2}} \right)^2$ is an increment of Brownian motion over a time interval of length $\frac{1}{2}\Delta t$, which is independent of $\mathcal{F}_{j+\frac{1}{2}}$. The result is

$$\mathrm{E}\left[\, U_j^2 \right] = \Delta t \, \mathrm{E}\left[\left(f_{j+\frac{1}{2}} - f_j \right)^2 \right] \,.$$

We apply the continuity assumption (2), taking the Δt there to be $\frac{1}{2}2^{-m}$ here. This leads to the bound (C is from (2), $\Delta t = 2^{-m}$ is the one here):

$$\mathrm{E}\left[\, U_j^2 \right] \leq \tfrac{1}{2} C \Delta t^2 \,.$$

This leads to the bound

$$\mathrm{E}\left[\, S^2 \right] \leq\leq \tfrac{1}{2} C \Delta t \sum_{t_j < t} \Delta t \leq \tfrac{1}{2} C \Delta t t \,.$$

This proves the inequality (3).

The Ito isometry formula may be derived using similar reasoning. The expected value of the square is a double sum with off diagonals having expected value zero. We compute the expected square of the approximations and see that the Ito isometry formula comes out in the limit. We evaluate the diagonal terms using conditional expectation, this time with respect to $\mathcal{F}_j$.

$$\mathrm{E}\left[\left(X_t^{(m)} \right)^2 \right] = \mathrm{E}\left[\left(\sum_{t_j < t} f_j \left(W_{j+1} - W_j \right) \right)^2 \right] \,.$$

The off diagonal terms have the form

$$\mathrm{E}[f_j \left(W_{j+1} - W_j \right) f_k \left(W_{k+1} - W_k \right)] \,.$$

If $t_j > t_k$, we condition on $\mathcal{F}_j$ and everything is known except $W_{j+1} - W_j$, which has conditional expectation equal to zero in $\mathcal{F}_j$ by the independent increments property. For the diagonal terms, we calculate

$$\begin{aligned}
\mathrm{E}\Big[f_j^2\,(W_{j+1}-W_j)^2\Big] &= \mathrm{E}\Big[\mathrm{E}\Big[f_j^2\,(W_{j+1}-W_j)^2 \mid \mathcal{F}_j\Big]\Big] \\
&= \mathrm{E}\Big[\mathrm{E}\Big[(W_{j+1}-W_j)^2 \mid \mathcal{F}_j\Big]\,f_j^2\Big] \\
&= \Delta t\,\mathrm{E}\big[f_j^2\big]
\end{aligned}$$

Therefore,

$$\mathrm{E}\Big[\Big(X_t^{(m)}\Big)^2\Big] = \sum_{t_j<t} \mathrm{E}\big[f_j^2\big]\,\Delta t\;.$$

When you take the limit $\Delta t \to 0$, this becomes

$$\mathrm{E}\Big[(X_t)^2\Big] = \int_0^t \mathrm{E}[f_s]\;ds\;. \tag{8}$$

This is the Ito isometry formula.

It may be stated informally as $\mathrm{E}[\,dW_s dW_{s'}] = 0$ if $s \neq s'$ and $\mathrm{E}\big[\,dW_s^2 \mid \mathcal{F}_s\big] = ds$. This expresses the idea that dW_s is in the future of s. Then you could reason as follows, using the trick that the square of a sum (an integral) is the double sum (integral) of the expected values:

$$\begin{aligned}
\mathrm{E}\Bigg[\bigg(\int_0^t f_s dW_s\bigg)^2\Bigg] &= \int_0^t\int_0^t \mathrm{E}[f_s f_{s'} dW_s dW_{s'}] \\
&= \int_0^t \mathrm{E}\Big[f_s^2\mathrm{E}\Big[(dW_s)^2 \mid \mathcal{F}_s\Big]\Big] \\
&= \int_0^t \mathrm{E}\big[f_s^2\big]\;ds\;.
\end{aligned}$$

If you look in Wikipedia, you may find a similar informal calculation using a delta function.

3 Ito's lemma

The informal Ito's lemma for Brownian motion is

$$df(W_t,t) = \partial_w f(W_t,t)dW_t + \tfrac{1}{2}\partial_w^2 f(W_t,t)dt + \partial_t f(W_t,t)dt\;. \tag{9}$$

The solid mathematical fact that this expression expresses is

$$f(W_t,t) - f(0,0) = \int_0^t \partial_s f(W_s,s)dW_s + \int_0^t \big(\tfrac{1}{2}\partial_w^2 f(W_s,s) + \partial_t f(W_s,s)\big)\;ds \tag{10}$$

The first integral on the right is an Ito integral and the second is an ordinary (Riemann) integral.

Two mathematical ideas enter into Ito's lemma. One is the Taylor series to second order in dW and first order in dt. This is because $(dW)^2$ is on the order of dt, so it is a "small" term but not a "tiny" one. The other idea is replacing $(dW)^2$ with its expected value $\mathrm{E}\left[\,(dW)^2\right] = dt$.

We now verify Ito's lemma in the integral form first approximating it using $\Delta t > 0$ and then taking the limit $\Delta t \to 0$. The calculations have several parts, but the parts all use ideas we have seen already. There is an upcoming conflict of notation. The f used in Ito's lemma is not the f used as the integrand in the Ito integral.

We start with

$$f(W_t, t) - f(0,0) = \sum_{t_j < t} \left[f(W_{j+1}, t_{j+1}) - f(W_j, t_j) \right] \ . \tag{11}$$

Then we use the Taylor approximation

$$f(W_{j+1}, t_{j+1}) - f(W_j, t_j) \approx \begin{array}{l} \partial_w f(W_j, t_j)\,(W_{j+1} - W_j) \\ + \frac{1}{2}\partial_w^2 f(W_j, t_j)\,(W_{j+1} - W_j)^2 \\ + \partial_t f(W_j, t_j)\Delta t \end{array} \ . \tag{12}$$

The error in a multi-variate Taylor approximation like this may be bounded using the "first neglected terms". These are the lowest order Taylor "pieces" not used. Here, those involve $\partial_w^3 f(W_j, t_j)$, and $\partial_w \partial_t f(W_j, t_j)$, and $\partial_t^2 f(W_j, t_j)$. We will assume that each of these is bounded. [Every theorem needs hypotheses, and applied mathematicians often "neglect" to say them.] Then there is a Taylor remainder theorem for multi-dimensional Taylor series which implies that the error is bounded in terms of the corresponding powers of ΔW and Δt:

$$C \left[|W_{j+1} - W_j|^3 + |W_{j+1} - W_j|\, \Delta t + \Delta t^2 \right] \ . \tag{13}$$

We now have two tasks. We must show that these error terms make "tiny" contributions, in the sense that the sum of them goes to zero as $\Delta t \to 0$. This is easy but involves one calculation we have not done yet. We also must show that the three terms we kept in the Taylor approximation of Δf, when added up, converge to the three integrals on the right of (10). This is where we justify replacing ΔW^2 with Δt.

We first show that the supposedly tiny terms (13) are actually tiny in the technical sense. For example

$$\sum_{t_j < t} |W_{j+1} - W_j|^3 \to 0 \ \text{ almost surely, as } \Delta t \to 0 \ .$$

The Borel Cantelli reasoning applies here. If $Q^{(m)} \geq 0$ is a sequence of of numbers, and

$$\sum_{m=1}^{\infty} Q^{(m)} < \infty$$

then $Q^{(m)} \to 0$ as $m \to \infty$. If $Q^{(m)}$ is a sequence of random variables and

$$\sum_{m=1}^{\infty} \mathrm{E}\Big[Q^{(m)} \Big] < \infty$$

then

$$\sum_{m=1}^{\infty} Q^{(m)} < \infty \; , \quad \text{almost surely,}$$

so

$$Q^{(m)} \to 0 \; , \quad \text{as } m \to \infty \; , \quad \text{almost surely.}$$

Now, take

$$Q^{(m)} = \sum_{t_j < t} \left| W_{j+1} - W_j \right|^3 \; .$$

To calculate the expectation, note that $W_{j+1} - W_j$ is a Gaussian random variable with mean zero and variance Δt. If Z is a standard normal (having $Z \sim \mathcal{N}(0,1)$), these random variables have the same distribution

$$W_{j+1} - W_j \sim \sqrt{\Delta t} Z \; .$$

Therefore the following expected values also are equal, and we calculate

$$\begin{aligned}
\mathrm{E}\Big[\left| W_{j+1} - W_j \right|^3 \Big] &= \mathrm{E}\Big[\left| \sqrt{\Delta t} Z \right|^3 \Big] \\
&= \Delta t^{\frac{3}{2}} \mathrm{E}\Big[\left| Z \right|^3 \Big] \\
&= \Delta t^{\frac{3}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left| z \right|^3 e^{-\frac{z^2}{2}} \, dz \\
&= \Delta t^{\frac{3}{2}} \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} z^3 e^{-\frac{z^2}{2}} \, dz \\
&= \Delta t^{\frac{3}{2}} \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} z^2 \left(z e^{-\frac{z^2}{2}} \right) \, dz \\
&= \Delta t^{\frac{3}{2}} \frac{4}{\sqrt{2\pi}} \int_{0}^{\infty} z e^{-\frac{z^2}{2}} \, dz \\
&= \Delta t^{\frac{3}{2}} \frac{4}{\sqrt{2\pi}} \; .
\end{aligned}$$

The value of the constant $\frac{4}{\sqrt{2\pi}}$ does not matter, only that it is a constant. The $\Delta t^{\frac{3}{3}}$ scaling is important. This scaling is "obvious" (we are pretty sure it's true) just from the scaling $\Delta W \sim \Delta t^{\frac{1}{2}}$. The calculations are here because $|\Delta W|^3$ is a higher power than Δt. For fat tailed distributions, the scaling of a higher power may be different than the scaling of a lower power. The calculation shows that a Gaussian is not fat tailed in this sense (or in any other sense).

Going back, we have (the fact that $C = \frac{4}{\sqrt{2\pi}}$ is irrelevant)

$$\begin{aligned}
\mathrm{E}\Big[\,Q^{(m)}\Big] &= \sum_{t_j<t} \mathrm{E}\Big[\,|W_{j+1} - W_j|^3\Big] \\
&= C \sum_{t_j<t} \Delta t^{\frac{3}{2}} \\
&= C\Delta t^{\frac{1}{2}} \sum_{t_j<t} \Delta t \\
&= C\left(\sqrt{2}\right)^{-m} t
\end{aligned}$$

This gives a geometric series with a finite sum. The same reasoning applies to the other tiny terms in (13), except that you don't need Borel Cantelli for Δt^2.

Back to the Taylor terms on the right of (12). The sum involving the first term is

$$\sum_{t_j<t} \partial_w f(W_j, t_j)\,(W_{j+1} - W_j) \ .$$

We just showed that in the limit $m \to \infty$, this converges to the Ito integral

$$\int_0^t \partial_w f(W_s, s)\, dW_s \ .$$

The sum involving the last term is

$$\sum_{t_j<t} \partial_t f(W_j, t_j)\Delta t \ .$$

Ordinary calculus shows that the limit as $m \to \infty$ is

$$\int_0^t \partial_t f(W_s, s)\, ds \ .$$

Here is the sum of the middle terms, adding and subtracting the expected values:

$$\sum_{t_j<t} \tfrac{1}{2}\partial_w^2 f(W_j, t_j)\,(W_{j+1} - W_j)^2 = \sum_{t_j<t} \tfrac{1}{2}\partial_w^2 f(W_j, t_j)\Delta t + \sum_{t_j<t} \tfrac{1}{2}\partial_w^2 f(W_j, t_j)\,\Big[(W_{j+1} - W_j)^2 - \Delta t\Big] \ .$$

The first sum on the right converges to the Riemann integral

$$\int_0^t \tfrac{1}{2}\partial_w^2 f(W_s, s)\, ds \ .$$

The second sum on the right converges to zero almost surely as $m \to \infty$ by a Borel Cantelli argument.